

JOHNSON GRANT

IN-61

7306

P-40

EXPERT SYSTEM VERIFICATION AND VALIDATION SURVEY

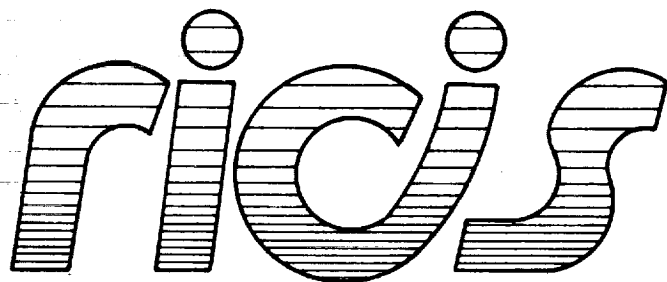
Delivery 2 - Survey Results

International Business Machines Corporation

August 15, 1990

**Cooperative Agreement NCC 9-16
Research Activity No. AI.16**

**NASA Johnson Space Center
Information Systems Directorate
Information Technology Division**



**Research Institute for Computing and Information Systems
University of Houston - Clear Lake**

N91-20789

Unclas

0007306

G3/61

(NASA-CR-188110) EXPERT SYSTEM VERIFICATION
AND VALIDATION SURVEY. DELIVERY 2: SURVEY
RESULTS (Houston Univ.) 40 p CSCI 09B

T · E · C · H · N · I · C · A · L R · E · P · O · R · T

The RICIS Concept

The University of Houston-Clear Lake established the Research Institute for Computing and Information systems in 1986 to encourage NASA Johnson Space Center and local industry to actively support research in the computing and information sciences. As part of this endeavor, UH-Clear Lake proposed a partnership with JSC to jointly define and manage an integrated program of research in advanced data processing technology needed for JSC's main missions, including administrative, engineering and science responsibilities. JSC agreed and entered into a three-year cooperative agreement with UH-Clear Lake beginning in May, 1986, to jointly plan and execute such research through RICIS. Additionally, under Cooperative Agreement NCC 9-16, computing and educational facilities are shared by the two institutions to conduct the research.

The mission of RICIS is to conduct, coordinate and disseminate research on computing and information systems among researchers, sponsors and users from UH-Clear Lake, NASA/JSC, and other research organizations. Within UH-Clear Lake, the mission is being implemented through interdisciplinary involvement of faculty and students from each of the four schools: Business, Education, Human Sciences and Humanities, and Natural and Applied Sciences.

Other research organizations are involved via the "gateway" concept. UH-Clear Lake establishes relationships with other universities and research organizations, having common research interests, to provide additional sources of expertise to conduct needed research.

A major role of RICIS is to find the best match of sponsors, researchers and research objectives to advance knowledge in the computing and information sciences. Working jointly with NASA/JSC, RICIS advises on research needs, recommends principals for conducting the research, provides technical and administrative support to coordinate the research, and integrates technical results into the cooperative goals of UH-Clear Lake and NASA/JSC.

***EXPERT SYSTEM VERIFICATION
AND
VALIDATION SURVEY***

Delivery 2 - Survey Results

Preface

This research was conducted under auspices of the Research Institute for Computing and Information Systems by the International Business Machines Corporation. Dr. Terry Feagin and Dr. T. F. Leibfried served as RICIS research representatives.

Funding has been provided by Information Technology Division, Information Systems Directorate, NASA/JSC through Cooperative Agreement NCC 9-16 between NASA Johnson Space Center and the University of Houston-Clear Lake. The NASA technical monitor for this activity was Chris Culbert, of the Software Technology Branch, Information Technology Division, Information Technology Directorate, NASA/JSC.

The views and conclusions contained in this report are those of the author and should not be interpreted as representative of the official policies, either express or implied, of NASA or the United States Government.

Expert System Verification and Validation Survey
RICIS Contract #069
Delivery 2 - Survey Results

August 15, 1990

IBM

3700 Bay Area Blvd.
Houston, Tx. 77058

Survey Results

Preface

This document constitutes the second delivery, "Survey Results," of the four deliveries scheduled for RICIS contract 069, "Verification and Validation of Expert Systems Study." The remaining two deliveries are:

1. Recommendations, due on August 30, 1990
2. Final Report, due on September 14, 1990

These final two deliveries will consist of updates to this document. "Recommendations" will be reported via a new section in this document following the section titled "Summary of Results." The "Final Report" will report survey data gathered late in the contract period via updates to the "Summary of Results," and may also include minor alterations to "Recommendations" based on this new data.

Note: The questionnaires actually used in the survey process are slightly different than those provided as the first delivery. The updated questionnaires are provided in Appendix A, "Expert Systems Evaluation Questionnaire (Developer)" on page 20 and Appendix B, "Expert Systems Evaluation Questionnaire (User)" on page 28.

Contents

Background	1
Survey Rationale	2
Purpose of the Questionnaires	3
Purpose of the Interviews	3
Survey Administration	4
Survey Questionnaires	5
Information Gathered	5
Human Factors	6
Summary of Results	7
General information	8
Performance Criteria	9
Requirements Definition	11
Development Information	12
V&V Activities Performed	14
V&V Issues Encountered	16
Recommendations	19
Appendix A. Expert Systems Evaluation Questionnaire (Developer)	20
Appendix B. Expert Systems Evaluation Questionnaire (User)	28

Background

The purpose of this task is to determine the state-of-the-practice in Verification and Validation (V&V) of Expert Systems (ESs) on current NASA and Industry applications. This is the first task of a series which has the ultimate purpose of ensuring that adequate ES V&V tools and techniques are available for Space Station Knowledge Based Systems development.

The strategy for determining the state-of-the-practice is to check how well each of the known ES V&V issues are being addressed and to what extent they have impacted the development of Expert Systems.

Note: This task does not attempt to prove or disprove whether Verification and Validation can or should be performed on Expert Systems. It is accepted that Verification and Validation should be applied to all software systems, including Expert Systems.

Survey Rationale

It is widely claimed that Expert Systems have been not been subject to the same level of Verification and Validation as traditionally developed software. Some people feel that this lack of V&V continues because of a "vicious circle," where nobody requires expert system V&V, so nobody does it. Consequently, since nobody knows how to do it, nobody requires it. There are two major reasons why the V&V process has not been documented: lack of a single life-cycle model, and technical differences between traditional software and expert systems.

Most expert system development life-cycles rely on iterative prototypes to develop the system behavior. This approach does not lead to methodical capture and documentation of the expected system behavior. Documented expectations, traditionally captured in a requirements document, are essential in the V&V process: you can't do testing if you don't know what to test for! One goal of this survey is to understand how the expected behavior of current expert systems is communicated and evaluated, even if a formal requirements document was not developed.

Expert Systems are typically composed of three parts: the knowledge base (KB), the inference engine, and the interface code between the inference engine and the peripheral devices (terminals, sensors, effectors, users, etc.). The inference engine and interface code are simply traditional software and should currently be V&Ved by accepted practices. This survey will help determine if these parts are V&Ved or whether, since they are part of an expert system, V&V is overlooked.

The knowledge base is the only part of the Expert System that raises new and unique issues. A set of of the *possible* issues are:

Issues primarily due to use of nonprocedural languages

- Understandability and readability to support inspections
- Testing coverage
- Standard validation tests for inference engines
- Real-time performance analysis

Issues due to heuristic knowledge (difficulty in organizing)

- Knowledge validation
- Modularity/Design

Issues primarily due to solving new complex problems

- Requirements
- Certification

Other issues

- Uncertainty Analysis
- Inheritance Process Test and Analysis
- Configuration Management

One of the purposes of this survey is to find out if these identified *possible* issues actually cause problems in practice, and if so, how the issues are being handled.

Purpose of the Questionnaires

Some of the information for this survey can be captured fairly easily and is accomplished through use of a questionnaire. The information captured this way includes:

- Application information - What kind of problem does the system address?, What are the performance goals?
- Expertise information - What was the relationship between the developers and expert(s)?, What is the performance level of the expert?
- Development information - How was the system developed?, How big is the system?
- Evaluation information - How was the system evaluated?
- Performance information - How important is good performance?, How well is the ES performing?

Purpose of the Interviews

The questionnaire answers lead to an additional set of questions involving the V&V issues described earlier. The additional questions are greatly affected by the answers provided in top questionnaire, so it would be more efficient to derive the information through direct interviews than to generate a large number of secondary questionnaires. The interviews attempt to uncover:

- the real issues involved in ES V&V (in comparison with the known *possible* issues outlined above).
- what is being done currently to address V&V (inspections, path testing, testing by the expert).
- what makes users trust the ESs, if the ESs are indeed trusted.
- what problems, unique to ESs, were encountered and possibly addressed during development and test.

The interviews are also required because we expect that some people will not fill out the questionnaires.

Survey Administration

This survey was designed so that the majority of the information would be gained from direct interviews with people involved in ES projects. Several people from each project, including developers, users, and managers, were interviewed to get a realistic view of the projects.

Several other activities were undertaken, both before and after the interview activity, to ensure that the results of the survey reflected the actual "state-of-the-practice". These activities included:

Identifying candidate ES projects

A list of projects to be contacted was created. The list included projects at NASA and IBM as well as projects from fields outside of the space industry.

Developing survey questionnaire(s)

To improve the chances of getting meaningful data from the questionnaire activity, separate questionnaires were developed for developers and users. Each questionnaire includes a question to indicate if the answers are from a manager or non-manager. Questionnaires are listed in Appendix A, "Expert Systems Evaluation Questionnaire (Developer)" on page 20 and Appendix B, "Expert Systems Evaluation Questionnaire (User)" on page 28.

Evaluating returned questionnaires

Each questionnaire was evaluated to determine if project interviews would uncover more information. If a project was to be interviewed, the questionnaire results provided guidance on which topics would be the most useful to explore.

Summarizing interview/questionnaire results

The summarized results of the questionnaire/interview activities are presented in section "Summary of Results" on page 7.

Recommendations

Recommendations for further action, based on the information in "Summary of Results" on page 7 will be provided as the next delivery.

Survey Questionnaires

Different versions of the questionnaire were developed for developers and users of the expert system. In addition, responses were expected to be different between managers and non-managers, so an indication is included on each questionnaire.

Information Gathered

Several types of information are captured by the questionnaire. Each question in the questionnaire addresses at least one of the previous types of information. For each type of information, the subtopics and questions which provide information are listed. The question numbers are noted as (development question, user question). Questions not available on a questionnaire are indicated by a "-".

General Information

Describes the general properties of the expert system, including the name (1, 41), a short description (4, 44), field of the problem (5, 45), and the type of problem to be solved (6, 46). Also captured are whether the survey taker was a manager (2, 42).

Performance Criteria

A major expertise issue is performance (probability that the results given are correct); specifically performance of the experts (10, 49), expected performance of the system (11, 50), and actual performance of the system (12, 51). Related to the performance issue is the amount of the problem space that the ES is expected to cover (8, 47), and that it actually covers (9, 48).

Requirements Definition

Requirements definition information includes how the requirements are documented (13, -), the difficulty in determining the requirements (14, -), and the availability of the expert(s) to resolve requirements issues during development (17, -). Influencing the performance issue is the number of experts (15, -), and whether the experts agree on the results obtained from the system (16, 61). It may also be useful to know if the expert (-, 52) and/or the developer(s) (18, 53) are part of the user organization.

Development Information

Development information that we are concerned with includes the development life-cycle used (19, -), and what languages and tools were used to develop the system (20, -). The size of the system (22, -), the total effort required for development, (29, -), and the effort required to develop the different parts of the ES (21, -) indicate the difficulty of the development effort. The sensitivity of the system (24, -) will influence the difficulty of future maintenance activities.

V&V Activities Performed

The major information to be captured during this task is the current state-of-the-practice for V&V of ESs, including the kinds of V&V being attempted, both during (28, -) and after (33, 60) development, and how much of the development effort was spent on V&V (30, -). Detailed information is also gathered for V&V activities for Knowledge Structures (25, -), the Inference Engine (26, -), and the Interface Code (27, -).

Information about the difficulty of the V&V effort (35, 62), whether a separate group performed V&V, (31, -) and how much effort was expended on the independent V&V (32, 59), is also gathered.

Whether the system is operational or prototype (3, 43), and the criticality of the system (37, 55) have an affect on the amount of V&V activities performed.

V&V Issues Encountered

If the state-of-the-practice is to be improved, the major issues that need to be addressed must be identified. One question (36, 63). directly asks whether each the known issues was actually encountered. Additional questions find out more information about specific issues, including the existence of certainty factors (7, -), whether configuration management was performed (34, -), and the difficulty of implementing the expertise through the Knowledge Structures (23, -). User acceptance is the ultimate test of the V&V activities. The comparison between expected system use (39, 57) and actual system use (40, 58), the perceived reliability of the system (38, 56), and why the user is convinced that the system produces correct results (-, 54) are all indicators of user acceptance.

Human Factors

The questionnaires were designed to capture as much accurate information as possible. In an effort to accomplish this, the following human factors issues were taken into account:

Questions should be understandable

Questions should have as few "technical" terms as possible to avoid confusion due to local usage. For questions that must have technical content, be sure to provide sufficient explanation.

Choices worded positively

Negatively worded choices may not get selected because the responder may feel there is something wrong with it.

Meaningful questions

The responder should feel that there is some purpose to the question.

Make use of fill-in-the-blank questions

The responder should not have to fill in long responses. Some questions can not have all possible responses enumerated, so the the user should be able to specify his own choice.

Summary of Results

The survey results are summarized in the following sections. The results are organized according to the type of information, as organized in "Information Gathered" on page 5. The numbers corresponding to the developer and user questionnaires, respectively, are given for each question. If the question is not in one of the questionnaires, the position is filled with a '-' (for example, if a question was number 10 in the developers questionnaire and not in the user questionnaire, the question numbers would be given as: 10, -). The total number of responses is also given for each question. The number of times each choice was selected is given to the left of the choice.

The following is a short summary of each type of information gathered.

General Information

Most of the respondents were involved with Expert Systems which perform Diagnosis (83%) in the Aerospace field (78%). The survey respondents were predominantly involved with development (84%).

Performance Criteria

The expected performance of the systems was nearly as high as the expert performance, but the actual performance was generally lower. The expected problem space coverage was not especially high; however, actual coverage was considerably less.

Requirements Definition

Of sixteen respondents, fourteen indicated that expert consultation was a basis for determining the behavior of the system. More revealing is that ten indicated consultation as the primary basis, while only eight indicated there was any written documentation. Seven respondents indicated that prototypes or similar tools were used for requirements.

Determining requirements had average difficulty. Availability of experts and agreement among experts were not problems.

Development Information

The most frequent Life-Cycle model used in the Cyclic Model (repetition of Requirements, Design, Rule Generation, and Prototyping until done). Most development was done with an Expert System shell (CLIPS, Knowledge Tool), and the predominant Interface Code was C and LISP. Applications were reasonably large and required an average of 38 person/months to develop. Developed systems were not reported to be particularly sensitive to change.

V&V Activities Performed

Most V&V activities relied on expert checking and comparison with expected results. Typically, 20% of the development effort was spent on V&V. The difficulty of the V&V effort was reported to be medium.

In most cases, there was not a separate group to perform V&V. When reported, the V&V effort expended varied widely between developers (1.7 person months) and users (22 person months). A possible explanation for the seemingly low amount of V&V performed is that 63% of the respondents indicated that the ES was a prototype system.

V&V Issues Encountered

The known issues most often cited as problems were: knowledge validation (62%), test coverage determination (50%), and problem complexity (50%). The least cited problem was analysis of certainty factors (only one respondent indicated that certainty factors were used). Every known issue was cited by at least one respondent.

The configuration management practices were reflective of the fact that a majority of the system were prototypes. The expected system use varied widely (3-2000), while actual system use was relatively good (less than half of the respondents provided information, suggesting that actual use was much lower than reported). System reliability, and expertise implementation difficulty were about average.

General information

The questions for the name of the ES, and the short description are not reported.

Field of the Problem

Question Numbers: 5, 45

Total Responses: 19

What field does the problem belong to?

- 15 Aerospace
- 2 Financial
- Information Systems
- Hardware
- Manufacturing
- Marketing
- Medical
- Personnel
- Research
- Service
- Software
- 2 Other

Type of Problem Solved

Question Numbers: 6, 46

Total Responses: 18

Which of the following items best describes the kind of problem the Expert System addresses? Please indicate primary purpose with a '*' and check all other applicable purposes (if any).

Note: The number of times the choice was selected as primary purpose is given in parentheses after the number of times the choice was selected.

- 3 (2) Design - Configuring objects under constraints
- 3 () Repair - Executing plans to administer prescribed remedies
- 3 (1) Control - Governing overall system behavior
- 4 () Planning - Designing actions
- 15 (9) Diagnosis - Inferring system malfunctions from observables
- 4 () Debugging - Prescribing remedies for malfunctions
- 8 () Prediction - Inferring likely consequences of given situations
- 8 (2) Monitoring - Comparing observations to expected outcomes

ORIGINAL PAGE IS
OF POOR QUALITY

- ☐ 1 () Instruction - Diagnosing, debugging, and repairing behavior
- ☐ 4 (2) Interpretation - Inferring situation descriptions from sensor data
- ☐ 2 (1) Classification - Categorizing objects by properties data

Role on Project

Question Numbers: 2, 42

Total Responses: 19

Were you a developer of the Expert System the manager of the, development organization, a user of the Expert System, or the manager of a department which uses the Expert System?

- ☐ 8 Developer of Expert System
- ☐ 3 Manager of Expert System development organization
- ☐ 5 Other Development
- ☐ 3 User of the Expert System
- ☐ Manager of a department using the Expert System
- ☐ Other User

Performance Criteria

Performance of the Experts

Question Numbers: 10, 49

Total Responses: 19

If human experts currently perform (or previously performed) the task, how often is the expert(s) expected to give the correct answer?

- ☐ Task not performed by human
- ☐ 4 "Correct" defined by expert
- ☐ 5 > 99%
- ☐ 5 95% to 99%
- ☐ 90% to 95%
- ☐ 1 80% to 90%
- ☐ 60% to 80%
- ☐ 1 40% to 60%
- ☐ 1 Other (100%)
- ☐ 2 I don't know

Expected Performance of the System

Question Numbers: 11, 50

Total Responses: 18

How often is the Expert System expected to provide the correct answer?

- ☐ 4 100%
- ☐ 6 > 99%
- ☐ 4 95% to 99%
- ☐ 3 90% to 95%
- ☐ 80% to 90%
- ☐ 60% to 80%
- ☐ 40% to 60%
- ☐ Other
- ☐ 1 I don't know

Actual Performance of the System

Question Numbers: 12, 51

Total Responses: 18

What is *your estimate* of how often the Expert System actually provides the correct answer?

- ☐ 1 100%
- ☐ 2 > 99%
- ☐ 3 95% to 99%
- ☐ 4 90% to 95%
- ☐ 1 80% to 90%
- ☐ 3 60% to 80%
- ☐ 1 40% to 60%
- ☐ 1 Other (< 40%)
- ☐ 2 I don't know

Expected Problem Space Coverage

Question Numbers: 8, 47

Total Responses: 19

How much of the problem space is the Expert System expected to cover?

- ☐ 4 100%
- ☐ 4 > 99%
- ☐ 2 95% to 99%
- ☐ 1 90% to 95%
- ☐ 2 80% to 90%
- ☐ 1 60% to 80%
- ☐ 1 40% to 60%
- ☐ 1 Other (25%)
- ☐ 2 I don't know

Actual Problem Space Coverage

Question Numbers: 9, 48

Total Responses: 16

What is *your estimate* of the problem space coverage actually provided by the Expert System?

- ☐ 2 100%
- ☐ > 99%
- ☐ 4 95% to 99%
- ☐ 90% to 95%
- ☐ 4 80% to 90%
- ☐ 4 60% to 80%
- ☐ 2 40% to 60%
- ☐ 2 Other (5%, < 40%)
- ☐ 3 I don't know

Requirements Definition

Requirements Format

Question Numbers: 13, -

Total Responses: 16

What was the basis for determining how the system was to behave? Please indicate the primary basis with a '*' and check all other applicable basis (if any).

Note: The number of times the choice was selected as primary basis is given in parentheses after the number of times the choice was selected.

- ☐ 2 () A pre-existing document _____
- ☐ 4 (1) A requirements document completed as part of development.
- ☐ 2 () Some other developed document _____
- ☐ 5 (1) A prototype of the system
- 14 (10) Expert consultation
- ☐ 3 () (user feedback, (2) similar tools)

Requirements Difficulty

Question Numbers: 14, -

Total Responses: 16

How difficult was it to develop the original concept of what the system was supposed to do?

- ☐ Trivial
- ☐ 4 Easy
- ☐ 8 Medium
- ☐ 4 Hard
- ☐ Impossible

Availability of the Expert(s)

Question Numbers: 17, -

Total Responses: 14

If the system was not developed by the expert, how much interaction was there between the expert(s) and the development team?

- ☐ 1 System was developed by expert
- ☐ 2 Constant
- ☐ 5 Frequent
- ☐ 3 Regular
- ☐ 3 Occasional
- ☐ None

Number of Experts

Question Numbers: 15, -

Total Responses: 16

Was more than one expert consulted during the development of the system?

- ☐ 2 System was developed by expert
- ☐ 2 Single expert
- ☐ 4 Multiple experts with lead
- ☐ 4 Committee of experts

_4 Other (no experts, experts as available, (2) multiple changing experts)

Agreement Among Experts

Question Numbers: 16, 61

Total Responses: 16

If more than one expert was available for consulting, how often did the experts agree on what results the Expert System was supposed to provide?

_3 A single expert was involved

_1 Always agree

_2 Agree 84% of the time (range 50%-99%)

Expert in User Organization

Question Numbers: -, 52

Total Responses: 3

Was the expert(s) a member of the user organization?

_3 Yes

— No

— User organization provided some expertise

Developers in User Organization

Question Numbers: 18, 53

Total Responses: 17

Was the developer(s) of the Expert System part of the user organization?

_8 Yes

_5 No

_4 Some development provided by user organization

Development Information

Development Life-Cycle Used

Question Numbers: 19, -

Total Responses: 16

Please indicate which development model was used for developing the Expert System.

_3 Requirements gathering preceded Design, Implementation, and Test (Traditional waterfall life-cycle).

_2 Requirements gathered before development of a prototype. A second requirements activity preceded Design, Implementation, and Test.

_7 Repetition of the Requirements, Design, Rule Generation, and Prototyping phases until production system (final prototype) was developed.

_3 No effort was made to follow a particular model.

_1 Other

Languages and Tools Used

Question Numbers: 20, -

Total Responses: 16

What was the primary language/tool for each part of the Expert System?

Note: The most frequent languages/tools are reported after the choice as: "frequency - language/tool."

13 Knowledge Structures (4 - LISP, 3 - CLIPS, others)

15 Inference Engine (4 - LISP, 3 - CLIPS, 3 - Knowledge Tool, others)

13 Interface Code (7 - C, 4 - LISP, others)

Size of the System

Question Numbers: 22, -

Total Responses: 16

Since Knowledge Bases can be written using several type of Knowledge Structures, please indicate how many of the following structures were used. If another type of structure was used, please describe it and how many were used.

Note: The number of times that a value was given for each choice is provided in parentheses following the number of times that the choice was selected. The range of the responses is given in parentheses after each choice.

14 (8) 100 Rules (range 30-300)

4 (1) 120 Frames

4 (2) 500 Facts (range 400-600)

4 (3) 68 Parameters (range 50-105)

1 (1) 35K Statements

3 (0) Other

Total Development Effort

Question Numbers: 29, -

Total Responses: 13

How much effort was expended in developing the system, including evaluation activities performed by the developers? 38 (range 3-200) person/months.

Detailed Development Effort

Question Numbers: 21, -

Total Responses: 15

What percentage of the total development effort was dedicated to each part of the Expert System?

Note: The number of times that a choice was selected is provided in parentheses before the average percentage of effort dedicated to the selected choice. The range of the responses is given in parentheses after each choice.

(15) 55 % Knowledge Structures (range 10%-100%)

(4) 10 % Inference Engine (range 10%-80%)

(14) 35 % Interface Code (range 10%-80%)

System Sensitivity

Question Numbers: 24, -

Total Responses: 16

When changes were made to the knowledge structures, how often did some unexpected result occur?

- ☐ 1 Never
- ☐ 11 Occasionally
- ☐ 3 Frequently
- ☐ 1 Usually
- ☐ Always

V&V Activities Performed

V&V Activities during development

Question Numbers: 28, -

Total Responses: 16

What testing activities were performed on the executing system? (indicate any that apply)

- ☐ 1 No evaluation was performed
- ☐ 11 Checked by expert(s)
- ☐ 12 Compared with expected results
- ☐ 6 Structural testing (e.g. cover all rules)
- ☐ 2 Other

V&V Activities after development

Question Numbers: 33, 60

Total Responses: 13

What testing activities were performed on the executing system before the system was delivered to the users? (indicate any that apply)

- ☐ 1 No evaluation was performed
- ☐ 9 Checked by expert(s)
- ☐ 10 Compared with expected results
- ☐ 7 User acceptance
- ☐ 4 System run in parallel
- ☐ 1 Other

Development effort was spent on V&V

Question Numbers: 30, -

Total Responses: 13

How much of the development effort was spent on evaluation? 20 % (range 0%-60%)

V&V of Knowledge Structures

Question Numbers: 25, -

Total Responses: 15

What evaluation activities were performed on the Knowledge Structures? (indicate any that apply)

- ☐ 1 No evaluation was performed
- ☐ 7 Desk checking
- ☐ 3 Formal inspections
- ☐ 8 Checked by expert(s)
- ☐ 7 Structural testing (e.g. cover all rules)
- ☐ 2 Other

V&V of Inference Engine

Question Numbers: 26, -

Total Responses: 14

What evaluation activities were performed on the Inference Engine? (indicate any that apply)

- ☐ 10 No evaluation was performed (ES shell was used)
- ☐ 2 No evaluation was performed
- ☐ Desk checking
- ☐ 1 Formal inspections
- ☐ 2 Structural testing
- ☐ 2 Other

V&V of Interface Code

Question Numbers: 27, -

Total Responses: 14

What evaluation activities were performed on the Interface Code? (indicate any that apply)

- ☐ 3 No evaluation was performed
- ☐ 7 Desk checking
- ☐ 2 Formal inspections
- ☐ 7 Structural testing (branch or path)
- ☐ 3 Other

Difficulty of V&V

Question Numbers: 35, 62

Total Responses: 13

Compared to conventional software testing efforts, how difficult was the evaluation of the Expert System?

- ☐ Trivial
- ☐ 3 Easy
- ☐ 6 Medium
- ☐ 4 Hard
- ☐ Impossible
- ☐ No evaluation was done

Separate V&V group

Question Numbers: 31, -
Total Responses: 15

Did a separate organization evaluate the Expert System before it was delivered to the users?

- _4 Yes, there was a separate evaluation organization.
- 11 No, there was not a separate evaluation organization.

Independent V&V Effort

Question Numbers: 32, 59
Total Responses: 4

If there was a separate evaluation team, how much effort was expended by the team in evaluating the correctness of the Expert System?

- (2) 1.7 (range .5-3) person/months reported by developers
- (2) 22 (range (20-24) person/months reported by users

Operational or Prototype System

Question Numbers: 3, 43
Total Responses: 19

Is the Expert System operational or is it a prototype?

- _6 Operational system
- 12 Prototype system
- _1 Operational prototype (write in)

System Criticality

Question Numbers: 37, 55
Total Responses: 19

How reliable is the Expert System required to be?

- _4 Trusted with human life
- _3 Trusted with mission objectives
- 10 As reliable as the expert
- _5 Assists the expert
- _7 Assists the user
- _ Other

V&V Issues Encountered

Known Issues Actually Encountered

Question Numbers: 36, 63
Total Responses: 16

Many people feel that some development issues are more of a problem with Expert Systems than with conventional systems. Which (if any) of the following were problems during implementation or test of this Expert System?

- _3 Understandability and readability of knowledge structures
- _8 Determining test coverage for knowledge structures
- _3 Modularity/Design of knowledge structures

Survey Results

- 10 Knowledge validation
- _1 Analysis of Certainty Factors
- _2 Validating the inference engine
- _5 Real-time performance analysis
- _8 Complexity of the Problem
- _5 Certification
- _1 Configuration Management
- _1 Other

Certainty Factors

Question Numbers: 7, -
Total Responses: 16

Does the Expert System include certainty factors?

- _1 Yes
- 13 No
- _2 I don't know

Configuration Management

Question Numbers: 34, -
Total Responses: 7

How were changes to the Expert System distributed to the users?

- _2 User updated system at developer's direction
- _1 Developers made changes to users' system
- _1 Untested system distributed to users
- _3 Tested system distributed to the users
- Configuration management group distributes system
- Other

Expertise Implementation Difficulty

Question Numbers: 23, -
Total Responses: 16

Aside from any difficulties in developing the original concept, how difficult was it to express the behavior (through the Knowledge Structures) of the expert?

- Trivial
- _2 Easy
- _7 Medium
- _7 Hard
- Impossible

Expected System Use

Question Numbers: 39, 57
Total Responses: 14

How many people are expected to make use of the Expert System? 252 (range 3-2000)

Actual System Use

Question Numbers: 40, 58

Total Responses: 8

How frequently are the (expected) users actually using the system? (Numbers may add up to more than 100% if the actual number of users is greater than the expected users.)

Note: The number of times a value was given is provided in parentheses before the percentage of use corresponding to each choice.

- (2) 3 % use the system more than expected (range 5%-15%)
- (7) 49 % use the system about as much as expected (range 10%-80%)
- (8) 26 % use the system less than expected (range 15%-90%)
- (5) 22 % do not use the system (range 10%-90%)

Perceived System Reliability

Question Numbers: 38, 56

Total Responses: 19

Does the Expert System seem to be more reliable or less reliable than conventional systems that are in use?

- _1 Significantly more reliable
- _4 More reliable
- _1 Slightly more reliable
- _3 Similar reliability
- _1 Slightly less reliable
- _1 Less reliable
- _1 Significantly less reliable
- _6 No comparison is available
- _3 I don't know

User Trust

Question Numbers: -, 54

Total Responses: 3

Why do you believe the results that the system gives?

- _1 Expert says it is correct
- _2 Participated in evaluation
- _1 Someone I trust did evaluation
- _3 Personal use and checking
- _1 User acceptance
- _1 I don't trust the results
- _1 Other

Recommendations

This section will be provided as the "Recommendations" delivery.

Appendix A. Expert Systems Evaluation Questionnaire (Developer)

By filling out this NASA funded questionnaire, you can help define the state-of-the-practice in the formal evaluation of Expert Systems on current NASA and industry applications. The information that you provide will be merged with the information from all other surveyed projects for the purpose of recommending future research and development activities. Individual responses are used solely as input to this information merging process. Each survey participant will be sent a copy of the final survey results.

Expert System applications are becoming more prevalent in fields where proper functioning is essential, such as the aerospace, medical, and financial industries. It is widely claimed that Expert Systems are not as rigorously evaluated as traditional software because of unique, unresolved evaluation issues. To ensure the continued and safe deployment of Expert Systems into critical areas, adequate evaluation techniques which address these issues must be developed and performed.

Instructions

The following questions concern your experiences with an Expert System, either as a developer or as the manager of the development effort. Feel free to indicate your answers in any way you like. Some of the choices on the multiple choice questions have places to fill in additional information; please indicate the choice and include the additional information, if possible. If you have any comments about the questions or your answers, please write them in the left margin.

Analysis of the responses may indicate that further discussion is required for complete understanding of the issues encountered during the evaluation process. Discussions will be held either as short one-on-one meetings or by telephone. Would you be available, at your convenience, to discuss the evaluation process in more detail?

Yes I am available for discussions.

Name _____

Phone _____

No I am not available for discussions.

If you have any questions regarding this questionnaire, please contact Keith Kelley at (713) 282-7303. If possible, please return completed questionnaires within one week of receipt to:

Keith Kelley
MC 6606
IBM Federal Sector Division
3700 Bay Area Blvd.
Houston, Tx. 77058-1199

Definitions

Certainty factors

Some problems require the use of certainty factors (also called probabilities, or fuzzy logic) in their processing. Facts which contain certainty factors have the form: "if a is true, then there is an x% chance that b is true."

Expert

The person who provides the knowledge that is to be captured in the Expert System.

Inference engine

Processes the knowledge structures to infer a set of output facts from a set of input facts. Examples of commercial systems are CLIPS and ESE.

Interface code

Used to supplement the inference process. Examples are interfacing the inference engine to a device, and performing arithmetic calculations.

Knowledge structures

Declarative part of the Expert System which represents the knowledge (typically called the Knowledge Base). Examples are frames and rules.

Problem space

The total number of cases which could potentially be addressed by the Expert System.

Problem space coverage

The percentage of the problem space that is addressed by the Expert System. For example, if the Expert System is supposed to be able to diagnose 100 malfunctions, but the total number of malfunctions is known to be 200, the problem space coverage is 50%.

Questions

1. What is the name of the Expert System you were/are involved with?

2. Were you a developer of the Expert System or the manager of the development organization?
 - a. Developer of Expert System
 - b. Manager of Expert System development organization
 - c. Other _____
3. Is the Expert System operational or is it a prototype?
 - a. Operational system
 - b. Prototype system
4. Briefly describe what the expert system does.

5. What field does the problem belong to?

a. Aerospace	g. Medical
b. Financial	h. Personnel
c. Information Systems	i. Research
d. Hardware	j. Service
e. Manufacturing	k. Software
f. Marketing	l. Other _____

6. Which of the following items best describes the kind of problem the Expert System addresses? Please indicate primary purpose with a '*' and check all other applicable purposes (if any).
 - a. Design - Configuring objects under constraints
 - b. Repair - Executing plans to administer prescribed remedies
 - c. Control - Governing overall system behavior
 - d. Planning - Designing actions
 - e. Diagnosis - Inferring system malfunctions from observables
 - f. Debugging - Prescribing remedies for malfunctions
 - g. Prediction - Inferring likely consequences of given situations
 - h. Monitoring - Comparing observations to expected outcomes
 - i. Instruction - Diagnosing, debugging, and repairing behavior
 - j. Interpretation - Inferring situation descriptions from sensor
 - k. Classification - Categorizing objects by properties data

7. Does the Expert System include certainty factors?

a. Yes	c. I don't know
b. No	

8. How much of the problem space is the Expert System expected to cover?

a. 100%	f. 60% to 80%
b. > 99%	g. 40% to 60%
c. 95% to 99%	h. Other _____%
d. 90% to 95%	i. I don't know
e. 80% to 90%	

9. What is *your estimate* of the problem space coverage actually provided by the Expert System?

a. Same as expected	f. 80% to 90%
b. 100%	g. 60% to 80%
c. > 99%	h. 40% to 60%
d. 95% to 99%	i. Other _____%
e. 90% to 95%	j. I don't know

Questions 10 through 12 are concerned with the percentage of problems within the problem space (covered by the Expert System) that are answered correctly.

10. If human experts currently perform (or previously performed) the task, how often is the expert(s) expected to give the correct answer?

a. Task not performed by human	f. 80% to 90%
b. "Correct" defined by expert	g. 60% to 80%
c. > 99%	h. 40% to 60%
d. 95% to 99%	i. Other _____ %
e. 90% to 95%	j. I don't know

11. How often is the Expert System expected to provide the correct answer?

a. 100%	f. 60% to 80%
b. > 99%	g. 40% to 60%
c. 95% to 99%	h. Other _____ %
d. 90% to 95%	i. I don't know
e. 80% to 90%	

12. What is *your estimate* of how often the Expert System *actually* provides the correct answer?

a. 100%	f. 60% to 80%
b. > 99%	g. 40% to 60%
c. 95% to 99%	h. Other _____ %
d. 90% to 95%	i. I don't know
e. 80% to 90%	

13. What was the basis for determining how the system was to behave? Please indicate the primary basis with a '*' and check all other applicable basis (if any).

a. A pre-existing document _____	
b. A requirements document completed as part of development.	
c. Some other developed document _____	
d. A prototype of the system	
e. Expert consultation	
f. Other _____	

14. How difficult was it to develop the original concept of what the system was supposed to do?

a. Trivial	d. Hard
b. Easy	e. Impossible
c. Medium	

15. Was more than one expert consulted during the development of the system?

a. System was developed by expert	c. Multiple experts with lead
b. Single expert	d. Committee of experts
	e. Other _____

16. If more than one expert was available for consulting, how often did the experts agree on what results the Expert System was supposed to provide?
 - a. A single expert was involved
 - b. Always agree
 - c. Agree _____% of the time.
17. If the system was not developed by the expert, how much interaction was there between the expert(s) and the development team?
 - a. System was developed by expert
 - b. Constant
 - c. Frequent
 - d. Regular
 - e. Occasional
 - f. None
18. Was the developer(s) part of the user organization?
 - a. Yes
 - b. No
 - c. Some developers were in the user organization
19. Please indicate which development model was used for developing the Expert System.
 - a. Requirements gathering preceded Design, Implementation, and Test (Traditional waterfall life-cycle).
 - b. Requirements gathered before development of a prototype. A second requirements activity preceded Design, Implementation, and Test.
 - c. Repetition of the Requirements, Design, Rule Generation, and Prototyping phases until production system (final prototype) was developed.
 - d. No effort was made to follow a particular model.
 - e. Other _____
20. What was the primary language/tool for each part of the Expert System?
 - a. Knowledge Structures _____
 - b. Inference Engine _____
 - c. Interface Code _____
21. What percentage of the total development effort was dedicated to each part of the Expert System?
 - a. Knowledge Structures _____%
 - b. Inference Engine _____% (If an Expert System Shell was used, this value should be 0%.)
 - c. Interface Code _____%

22. Since Knowledge Bases can be written using several type of Knowledge Structures, please indicate how many of the following structures were used. If another type of structure was used, please describe it and how many were used.

- | | |
|-----------------|-----------------------------|
| a. Rules _____ | d. Parameters _____ |
| b. Frames _____ | e. Statements _____ |
| c. Facts _____ | f. Other (#) _____ of _____ |

23. Aside from any difficulties in developing the original concept, how difficult was it to express the behavior (through the Knowledge Structures) of the expert?

- | | |
|------------|---------------|
| a. Trivial | d. Hard |
| b. Easy | e. Impossible |
| c. Medium | |

24. When changes were made to the knowledge structures, how often did some unexpected result occur?

- | | |
|-----------------|------------|
| a. Never | d. Usually |
| b. Occasionally | e. Always |
| c. Frequently | |

Questions 25 through 28 are concerned with the evaluation activities performed during development.

25. What evaluation activities were performed on the knowledge Structures? (indicate any that apply)

- | | |
|--------------------------------|--|
| a. No evaluation was performed | d. Checked by expert(s) |
| b. Desk checking | e. Structural testing (e.g. cover all rules) |
| c. Formal inspections | f. Other _____ |

26. What evaluation activities were performed on the Inference Engine? (indicate any that apply)

- | | |
|--------------------------------|-----------------------|
| a. No evaluation was performed | d. Structural testing |
| b. Desk checking | e. Other _____ |
| c. Formal inspections | |

27. What evaluation activities were performed on the Interface Code? (indicate any that apply)

- | | |
|--------------------------------|--|
| a. No evaluation was performed | d. Structural testing (branch or path) |
| b. Desk checking | e. Other _____ |
| c. Formal inspections | |

28. What testing activities were performed on the executing system? (indicate any that apply)
- | | |
|-----------------------------------|--|
| a. No evaluation was performed | d. Structural testing (e.g. cover all rules) |
| b. Checked by expert(s) | e. Other _____ |
| c. Compared with expected results | |
29. How much effort was expended in developing the system, including evaluation activities performed by the developers? _____ person/months.
30. How much of the development effort was spent on evaluation? _____ %.
31. Did a separate organization evaluate the Expert System before it was delivered to the users?
- | | |
|---|--|
| a. Yes, there was a separate evaluation organization. | b. No, there was not a separate evaluation organization. |
|---|--|
32. If there was a separate evaluation team, how much effort was expended by the team in evaluating the correctness of the Expert System? _____ person/months.
33. What testing activities were performed on the executing system before the system was delivered to the users? (indicate any that apply)
- | | |
|-----------------------------------|---------------------------|
| a. No evaluation was performed | d. User acceptance |
| b. Checked by expert(s) | e. System run in parallel |
| c. Compared with expected results | f. Other _____ |
34. How were changes to the Expert System distributed to the users?
- | |
|--|
| a. User updated system at developer's direction |
| b. Developers made changes to users' system |
| c. Untested system distributed to users |
| d. Tested system distributed to the users |
| e. Configuration management group distributes system |
| f. Other _____ |
35. Compared to conventional software testing efforts, how difficult was the evaluation of the Expert System?
- | | |
|------------|---------------------------|
| a. Trivial | d. Hard |
| b. Easy | e. Impossible |
| c. Medium | f. No evaluation was done |

36. Many people feel that some development issues are more of a problem with Expert Systems than with conventional systems. Which (if any) of the following were problems during implementation or test of this Expert System?
- a. Understandability and readability of knowledge structures
 - b. Determining test coverage for knowledge structures
 - c. Modularity/Design of knowledge structures
 - d. Knowledge validation
 - e. Analysis of Certainty Factors
 - f. Validating the inference engine
 - g. Real-time performance analysis
 - h. Complexity of the Problem
 - i. Certification
 - j. Configuration Management
 - k. Other _____
37. How reliable is the Expert System required to be?
- a. Trusted with human life
 - b. Trusted with mission objectives
 - c. As reliable as the expert
 - d. Assists the expert
 - e. Assists the user
 - f. Other _____
38. Does the Expert System seem to be more reliable or less reliable than conventional systems that are in use?
- a. Significantly more reliable
 - b. More reliable
 - c. Slightly more reliable
 - d. Similar reliability
 - e. Slightly less reliable
 - f. Less reliable
 - g. Significantly less reliable
 - h. No comparison is available
 - i. I don't know
39. How many people are expected to make use of the Expert System?

40. How frequently are the (expected) users actually using the system? (Numbers may add up to more than 100% if the actual number of users is greater than the expected users.)
- a. _____% use the system more than expected
 - b. _____% use the system about as much as expected
 - c. _____% use the system less than expected
 - d. _____% do not use the system

Appendix B. Expert Systems Evaluation Questionnaire (User)

By filling out this NASA funded questionnaire, you can help define the state-of-the-practice in the formal evaluation of Expert Systems on current NASA and industry applications. The information that you provide will be merged with the information from all other surveyed projects for the purpose of recommending future research and development activities. Individual responses are used solely as input to this information merging process. Each survey participant will be sent a copy of the final survey results.

Expert System applications are becoming more prevalent in fields where proper functioning is essential, such as the aerospace, medical, and financial industries. It is widely claimed that Expert Systems are not as rigorously evaluated as traditional software because of unique, unresolved evaluation issues. To ensure the continued and safe deployment of Expert Systems into critical areas, adequate evaluation techniques which address these issues must be developed and performed.

Instructions

The following questions concern your experiences with an Expert System, either as a user or as the manager of a department that uses Expert System. Feel free to indicate your answers in any way you like. Some of the choices on the multiple choice questions have places to fill in additional information; please indicate the choice and include the additional information, if possible. If you have any comments about the questions or your answers, please write them in the left margin.

Analysis of the responses may indicate that further discussion is required for complete understanding of the issues encountered during the evaluation process. Discussions will be held either as short one-on-one meetings or by telephone. Would you be available, at your convenience, to discuss the evaluation process in more detail?

Yes I am available for discussions.

Name _____

Phone _____

No I am not available for discussions.

If you have any questions regarding this questionnaire, please contact Keith Kelley at (713) 282-7303. If possible, please return completed questionnaires within one week of receipt to:

Keith Kelley
MC 6606
IBM Federal Sector Division
3700 Bay Area Blvd.
Houston, Tx. 77058-1199

Definitions

Expert

The person who provides the knowledge that is to be captured in the Expert System.

Inference engine

Processes the knowledge structures to infer a set of output facts from a set of input facts. Examples of commercial systems are CLIPS and ESE.

Knowledge structures

Declarative part of the Expert System which represents the knowledge (typically called the Knowledge Base). Examples are frames and rules.

Problem space

The total number of cases which could potentially be addressed by the Expert System.

Problem space coverage

The percentage of the problem space that is addressed by the Expert System. For example, if the Expert System is supposed to be able to diagnose 100 malfunctions, but the total number of malfunctions is known to be 200, the problem space coverage is 50%.

Questions

41. What is the name of the Expert System you were/are involved with?

42. Are you a user of the Expert System or the manager of a department which uses the Expert System?
- a. User of the Expert System
 - b. Manager of a department using the Expert System
 - c. Other _____
43. Is the Expert System operational or is it a prototype?
- a. Operational system
 - b. Prototype system
44. Briefly describe what the expert system does.

45. What field does the problem belong to?
- | | |
|------------------------|----------------|
| a. Aerospace | g. Medical |
| b. Financial | h. Personnel |
| c. Information Systems | i. Research |
| d. Hardware | j. Service |
| e. Manufacturing | k. Software |
| f. Marketing | l. Other _____ |

46. Which of the following items best describes the kind of problem the Expert System addresses? Please indicate primary purpose with a "*" and check all other applicable purposes (if any).

- a. Design - Configuring objects under constraints
- b. Repair - Executing plans to administer prescribed remedies
- c. Control - Governing overall system behavior
- d. Planning - Designing actions
- e. Diagnosis - Inferring system malfunctions from observables
- f. Debugging - Prescribing remedies for malfunctions
- g. Prediction - Inferring likely consequences of given situations
- h. Monitoring - Comparing observations to expected outcomes
- i. Instruction - Diagnosing, debugging, and repairing behavior
- j. Interpretation - Inferring situation descriptions from sensor data
- k. Classification - Categorizing objects by properties

47. How much of the problem space is the Expert System **expected** to cover?

- a. 100%
- b. > 99%
- c. 95% to 99%
- d. 90% to 95%
- e. 80% to 90%
- f. 60% to 80%
- g. 40% to 60%
- h. Other _____ %
- i. I don't know

48. What is *your estimate* of the problem space coverage **actually** provided by the Expert System?

- a. Same as expected
- b. 100%
- c. > 99%
- d. 95% to 99%
- e. 90% to 95%
- f. 80% to 90%
- g. 60% to 80%
- h. 40% to 60%
- i. Other _____ %
- j. I don't know

Questions 49 through 51 are concerned with the percentage of problems within the problem space (covered by the Expert System) that are answered correctly.

49. If human experts currently perform (or previously performed) the task, how often is the expert(s) expected to give the correct answer?

- a. Task not performed by human
- b. "Correct" defined by expert
- c. > 99%
- d. 95% to 99%
- e. 90% to 95%
- f. 80% to 90%
- g. 60% to 80%
- h. 40% to 60%
- i. Other _____ %
- j. I don't know

50. How often is the Expert System expected to provide the correct answer?

- a. 100%
- b. > 99%
- c. 95% to 99%
- d. 90% to 95%
- e. 80% to 90%
- f. 60% to 80%
- g. 40% to 60%
- h. Other _____ %
- i. I don't know

51. What is *your estimate* of how often the Expert System actually provides the correct answer?
- | | |
|---------------|------------------|
| a. 100% | f. 60% to 80% |
| b. > 99% | g. 40% to 60% |
| c. 95% to 99% | h. Other _____ % |
| d. 90% to 95% | i. I don't know |
| e. 80% to 90% | |
52. Was the expert(s) a member of the user organization?
- | | |
|--------|--|
| a. Yes | c. User organization provided some expertise |
| b. No | |
53. Was the developer(s) of the Expert System part of the user organization?
- | | |
|--------|---|
| a. Yes | c. Some development provided by user organization |
| b. No | |
54. Why do you believe the results that the system gives?
- | | |
|-----------------------------------|------------------------------|
| a. Expert says it is correct | e. User acceptance |
| b. Participated in evaluation | f. I don't trust the results |
| c. Someone I trust did evaluation | g. Other _____ |
| d. Personal use and checking | |
55. How reliable is the Expert System required to be?
- | | |
|------------------------------------|-----------------------|
| a. Trusted with human life | d. Assists the expert |
| b. Trusted with mission objectives | e. Assists the user |
| c. As reliable as the expert | f. Other _____ |
56. Does the Expert System seem to be more reliable or less reliable than conventional systems that are in use?
- | | |
|--------------------------------|--------------------------------|
| a. Significantly more reliable | f. Less reliable |
| b. More reliable | g. Significantly less reliable |
| c. Slightly more reliable | h. No comparison is available |
| d. Similar reliability | i. I don't know |
| e. Slightly less reliable | |
57. How many people are expected to make use of the Expert System?

58. How frequently are the (expected) users actually using the system? (Numbers may add up to more than 100% if the actual number of users is greater than the expected users.)
- a. _____ % use the system more than expected
 - b. _____ % use the system about as much as expected
 - c. _____ % use the system less than expected
 - d. _____ % do not use the system

If you were not involved with evaluating the Expert System, please leave the remaining questions unanswered.

59. How much effort was expended by the evaluation team in evaluating the correctness of the Expert System? _____ person/months.
60. What testing activities were performed on the executing system before the system was delivered to the users? (indicate any that apply)
- a. No evaluation was performed
 - b. Checked by expert(s)
 - c. Compared with expected results
 - d. User acceptance
 - e. System run in parallel
 - f. Other _____
61. If more than one expert was available for consulting, how often did the experts agree on what results the Expert System is supposed to provide?
- a. No expert was involved
 - b. A single expert was involved
 - c. Always agree
 - d. Agree _____ % of the time.
62. Compared to conventional software testing efforts, how difficult was the evaluation of the Expert System?
- a. Trivial
 - b. Easy
 - c. Medium
 - d. Hard
 - e. Impossible
63. Many people feel that some development issues are more of a problem with Expert Systems than with conventional systems. Which (if any) of the following were problems during testing of the Expert System?
- a. Understandability and readability of knowledge structures
 - b. Determining test coverage for knowledge structures
 - c. Modularity/Design of knowledge structures
 - d. Knowledge validation
 - e. Analysis of Certainty Factors
 - f. Validating the inference engines
 - g. Real-time performance analysis
 - h. Complexity of the Problem
 - i. Certification
 - j. Other _____